The Journal of Wildlife Management 9999:1–8; 2012; DOI: 10.1002/jwmg.373

Research Article

Citizen Science and Observer Variability During American Pika Surveys

LUCAS MOYER-HORNER,[1] *Department of Zoology, University of Wisconsin, Madison, 250 N. Mills Street, WI 53706, USA*
MATTHEW M. SMITH, *Department of Zoology, University of Wisconsin, Madison, 250 N. Mills Street, WI 53706, USA*
JAMI BELT, *Crown of the Continent Research Learning Center, Glacier National Park, West Glacier, MT 59936, USA*

ABSTRACT Within- and between-group observer variability can confound scientific discovery. If observer variability can be quantified and is addressed, data collected by participants with wide ranges of experience and training can yield more reliable inferences. The American pika (*Ochotona princeps*) is a mammalian sentinel of climate change that has received consideration for listing under the United States Endangered Species Act. As a result, numerous pika monitoring initiatives have been started throughout the mountains in western North America. Some initiatives employ research teams of biological science technicians (professionals), whereas many rely on networks of citizen scientists, or volunteers, for data collection. To date, few studies have quantified observer variability during pika surveys; none have explored the reliability of professional crews or volunteers. We conducted pika surveys in Glacier National Park, Montana, to quantify observer variability. We investigated observer variability 1) among a crew of professionals, 2) among volunteers, and 3) between professionals and volunteers. Professionals were more consistent at identifying pika signs and estimating potential home ranges and consistently found more pika signs than did the volunteers, with the exception of pika sightings. Estimates of pika occupancy were consistent at each site among volunteers conducting sitting surveys. We suggest that sitting surveys conducted by volunteers can reliably detect pika site occupancy. However, data on population dynamics of pikas (e.g., density) should be collected by professionals. Observer variability analyses of this nature should be common practice for wildlife-resource managers and scientists, especially with observers of varying levels of experience and motivation. © 2012 The Wildlife Society.

KEY WORDS American pika, citizen science, distribution, Glacier National Park, observer variability, *Ochotona princeps*, professional, volunteer.

Observer variability during wildlife population surveys confounds detection of change and treatment effects over time. Wildlife biologists are increasingly concerned with species' detectability, especially for secretive and low-density populations. When observer variability occurs, survey precision decreases and results can become inconsistent. This limits detection of trends and reduces confidence in data interpretation (MacKenzie et al. 2006). Observer variability has been shown to be prominent in wildlife-resource studies involving many species, such as frogs (e.g., Lotz and Allen 2007), birds (e.g., Diefenbach et al. 2003), and mammals (e.g., Newman et al. 2003). Observer variability can arise from a number of factors including: experience, environment, terrain, training, and motivation. Diefenbach et al. (2003) reported considerable variation among observers' ability to detect birds and found that experience influenced accuracy of species identification. In this case, variation arose from difficulties identifying closely related bird species by call recognition. During anuran call surveys, Lotz and Allen (2007) found inaccurate species identification because of background noise and poor recording devices; but did not find an experience-based difference. Spaulding et al. (2000) examined observer variability in identifying diet of gray wolves from scat. In this case, they found significant differences based on how thoroughly the observer dissected the scat. Observer variability may arise from variation in levels of both expertise and motivation. It can be exacerbated when monitoring is conducted across large, complex landscapes where detection is more challenging.

Managers responding to shortfalls in funding for monitoring are increasingly turning toward the use of citizen science programs (Silvertown 2009). Citizen science programs use trained citizen scientists to collect data without direct supervision (Trumbull et al. 2000). The use of volunteer-based programs can enable collection of large data sets at reduced costs (Danielsen et al. 2005, Newman et al. 2003). Citizen science offers many benefits for wildlife-resource monitoring, including increasing the scientific literacy of participants (Trumbull et al. 2000, Bonney et al. 2009) and building a sense of community among citizen scientists and land managers (e.g., National Park Service; Overdevest et al. 2004). In addition, the cost savings from using citizen scientists instead of paid professionals can increase project duration and sustainability (Danielsen et al. 2005).

Received: 28 June 2011; Accepted: 19 January 2012

---

[1] *E-mail: lrmoyerh@wisc.edu*

Proponents believe citizen science programs can be a reliable supplement to professional monitoring (Danielsen et al. 2005). However, managers are not always clear whether citizen scientists provide scientifically useful data (Silvertown 2009). Because citizen scientists (hereafter volunteers) often have a wide range of prior experience with ecological research, the potential for observer variability may be greater among volunteers than among professionals (Cohn 2008). Although managers must evaluate the reliability of data collected by professionals, they may need to scrutinize data collected by volunteers even more closely (Newman et al. 2003).

Mountain ecosystems encompass many of our parks and protected areas (Scott et al. 2001) and contain abundant wildlife resources, but are often challenging for wildlife-resource managers to monitor. Mountainous regions are also of particular concern because many are expected to experience some of the largest climatic changes in the coming decades (Root et al. 2003, Inouye 2008). Increased monitoring of particularly sensitive and detectable species in mountain landscapes is especially important for identifying trends, predicting effects, and making management decisions in these wildlife-rich regions. Because the vast landscapes and difficult terrain pose practical and financial difficulties for monitoring initiatives, managers often elect to monitor certain indicator species that may provide an early warning for changes in biodiversity and the environment (Lindenmayer and Likens 2011). Concern about the impacts of climate change on mountain-dwelling species and mountainous ecosystems has prompted managers to look for alpine indicator species (Heller and Zavaleta 2009).

Monitoring of American pikas (*Ochotona princeps*) may provide an opportunity to quantify the effects of climate change on alpine species. The American pika is highly sensitive to changes in temperature and may serve as a mammalian sentinel of climate change (Beever et al. 2003, 2010). Pikas are habitat specialists found only in talus deposits and talus-like formations, such as lava flows or mining ore dumps. This limits pika dispersal and can inhibit their ability to shift their range, in latitude or elevation, when temperatures rise (Beever et al. 2010).

In addition, the American pika is an ideal candidate species for a comparison of data collected by volunteers and professionals, because of their high detectability. Detectability estimates range from 80% to 95% (Beever et al. 2003, 2010; Ray and Beever 2007; Rodhouse et al. 2010). Pikas are relatively easy to find because of their diurnal activity, obligate relationship to talus, recognizable vocalizations, prominent food caches (hay piles), and identifiable scat (Smith 1974, Beever 2003). The high detectability of pikas, along with increasing concern about their sensitivity to climate change, has prompted the creation of numerous pika monitoring initiatives that employ professionals and volunteers. To date, however, no published studies have explored observer variability during monitoring of American pikas by professionals or volunteers.

The United States Fish and Wildlife Service (USFWS) recently completed a review of American pikas to determine whether the species warranted protection under the Endangered Species Act. The petition cited retraction of suitable habitat, caused by climate change, as the primary threat to American pikas. The USFWS denied listing, in part, because of insufficient baseline distribution and monitoring data (USFWS 2010). The review prompted many federal and state agency managers in the western United States to develop monitoring initiatives to address the need for baseline information on this easily detectable alpine species. Some projects have secured sufficient funding to enable the use of professional biological technicians (hereafter professionals) to collect data on pikas. However, without the funding that may have followed Endangered Species status, many managers are faced with inadequate resources to monitor American pikas using these traditional approaches.

The possible effect of observer variability is an important aspect in any observational study that relies on data gathered from multiple sources, whether professional or volunteer. With the growing number of initiatives focused on detecting population changes in American pikas over time, quantifying observer variability from professionals and volunteers is critical. We analyzed data from 2 monitoring programs that were focused on estimating baseline populations for American pikas in Glacier National Park, Montana. Our objective was to quantify and compare observer variability from professionals and volunteers on baseline and trend estimates of American pika populations and identify potential effects and causes. The professionals were not necessarily paid, but were full-time technicians on a field crew. In contrast, the volunteers received training, but conducted surveys on an occasional basis and without direct supervision.

We hypothesized that the high detectability of pikas would lead to insignificant observer variability among professionals, and between professionals and volunteers. We conducted 3 separate variability trials. First, professionals conducted pika detectability surveys to assess observer variability among the professionals. Second, we evaluated site surveys conducted separately by the professionals and volunteers to quantify observer variability between the groups. Finally, we evaluated sitting-surveys conducted by volunteers to quantify variability among volunteers.

STUDY AREA

We conducted our study during 2008 and 2009 in Glacier National Park, Montana. Glacier National Park contains 4,081 km^2 of federally protected land and is situated in the northern Rocky Mountains, Montana, USA. Over one-third of the park is within the alpine zone. Talus formations, which provide refuge for American pikas from thermal extremes and predation (Tyser 1980), are abundant in the area between 1,500 m and 2,000 m elevation. Glacier National Park contains >300 distinct talus patches, each ranging in size from 0.1 ha to 100 ha. Talus is typically found at the base of fractured cliffs and is generally discontinuous and localized in Glacier National Park, limiting pika dispersal. Surveyed sites occurred in alpine, sub-alpine, and mixed-conifer forest habitats. Sites consisted of discrete talus deposits containing rocks >0.2 m in diameter, and fringed by grasses and forbs.

METHODS

We collected data from 2 concurrent pika monitoring programs to assess pika detectability and to compare observer variability between and among professionals and volunteers. The professionals were employed for ≥4 weeks with a comprehensive pika monitoring project in Glacier National Park. Professionals received at least 2 days of training, including at least 2 practice surveys, and conducted 10–30 surveys/week. The volunteers were visitors to the park who volunteered on their own schedule without direct supervision for 1–10 surveys per season (Jul–Sep). Volunteers were required to participate in an 8-hour training period to learn how to monitor pikas, mountain goats (*Oreamnos americanus*), and Clark's nutcrackers (*Nucifraga columbiana*); 3 species that were being monitored as a part of Glacier National Park's High Country Citizen Science Program.

Pika Detectability Surveys Conducted by Professionals

To explore observer variability among professionals and quantify the variables that influenced detection of pika signs, we asked 10 trained professionals to individually survey 8 separate talus sites (section) near Snyder Lake, 12–13 July 2009 ($n = 80$ observations). The 8 sections occurred at 2 separate talus patches and were delineated by rock cairns. We showed each professional a map of the sections and their boundaries. We instructed professionals to spend approximately 30 minutes surveying each section. They recorded only signs that occurred within their current section. For example, they did not record calls (characteristic pika vocalizations) and sightings (observations of a physical pika) perceived to be beyond the boundary of their present section. Pikas could have moved between some of the sections between surveys. Additionally, estimation of caller location can be problematic for some species (e.g., Alldredge et al. 2008). However, these factors were minimized because: 1) trials occurred over a short time span, 2) pikas are highly philopatric (Smith and Weston 1990), 3) pikas occur in relatively low densities, and 4) pikas defend and move within territories much smaller in area than the sites themselves. We categorized surveys by time of day according to the survey's start time: morning = before 1200, midday = 1201–1700, and evening = after 1700. We assigned an experience level to professionals from 1–3 based on how many previous surveys they had conducted: less than 11 surveys (1), between 11 and 30 surveys (2), and more than 30 surveys (3). We also recorded the name of the professional conducting each survey to determine variability among individual observers.

Professionals searched the designated section for transient and static pika signs. Transient signs included calls and sightings as described above and indicated that pikas were currently present. Static signs indicated current or historic presence, and included scat and hay piles, both of which were easily identifiable. Each pika fecal pellet is spherical, and has a diameter of approximately 2.25–3.50 mm. Hay piles are a pika's collections of local vegetation that also contain some scat. Professionals also estimated visible hay-pile volume, length, width, and depth. Pikas are solitary central-place foragers, thus signs within a certain radius of a hay pile typically belong to a single individual. This territory radius can vary from 10 m to 25 m depending on a number of factors, including habitat quality, pika density, and geographic region (Smith and Weston 1990). In Glacier National Park, we observed that 1 potential pika home range was typically a 20-m radius circle (L. Moyer-Horner, University of Wisconsin, unpublished data). If observers found pika signs in an area without a visible hay pile, professionals assumed the hay pile was either deep in the talus or the site was newly occupied. Professionals estimated the center of a home range in those instances when a hay pile was not evident by the positions of other signs not assigned to another home range. They next estimated the number of potential home ranges in each section and the number that contained at least 1 pika sign (occupied home ranges). Professionals identified home ranges without pika signs by the presence of rocks >0.2 m in diameter. Detection of a sighting or call in any of the occupied home ranges indicated that 1 territorial pika was currently occupying that home range. Professionals thus recorded a current occupant when an occupied home range included at least 1 of these transient signs. For example, a professional could estimate that a site contained 20 home ranges, find 15 occupied home ranges, and hear or see a pika in 6 of the occupied home ranges (current occupants = 6).

Site Surveys Conducted by Volunteers and Professionals

To compare observer variability between professionals and volunteers in detection of static and transient pika signs, teams of professionals or volunteers surveyed 12 talus sites during 2008 and 2009 ($n = 22$). One or more surveyors navigated to the site using a provided site description and Global Positioning System (GPS) coordinates. Upon reaching the position, surveyors recorded a survey start time that was categorized by time of day as either morning, midday, or evening in the manner described above. Surveyors traversed the site by searching non-overlapping sections of the talus >20 m from one another, recording all the observed pika signs, and estimating occupied home ranges as described above.

Although the survey protocol followed by professionals and volunteers was the same, the length of each survey and the number of surveyors varied. To adjust for these differences we quantified survey effort as the number of active surveyors multiplied by the total search time (in minutes). Professionals always surveyed in pairs so their survey effort was the total search time multiplied by 2. Volunteer group-size varied from 1 to 4. When surveys were conducted in both 2008 and 2009, we identified them by both the site and the year, thus deriving a site × year variable. This allowed us to compare only surveys of the same site during the same year, thereby removing the effect of annual turnover in pika populations from analyses. For the analyses, observer type meant either professional or volunteer.

To represent seasonality, which may affect the level of pika activity, we calculated the number of days from 1 June to the survey date (season; e.g., a survey conducted on 11 Aug would have season = 71). In order to compare sites of different size and containing different numbers of pikas, we divided calls and sightings by occupied home ranges and current occupants.

Sitting Surveys Conducted by Volunteers

Volunteers conducted sitting surveys at 21 sites during 2008 and 2009 ($n = 95$ surveys). During sitting surveys, an individual volunteer sat on a rock that provided a good vantage point near that site's designated GPS position. They recorded time of day, sightings, calls, and estimated number of current occupants. Sitting surveys are ideal for volunteers because they reduce the safety risks and physical challenges associated with traversing talus, and thus can be conducted by volunteers with varying physical abilities. Additionally the simplified protocol minimizes the risk of surveyor error. Newman et al. (2003) found that techniques that did not require lengthy or specialized training during mammal monitoring produced reliable data.

Data Analysis

To describe variability in detection of pika signs between and among professionals and volunteers, we fit a series of linear regression models that included potential predictor variables chosen as plausible competing hypotheses. Potential predictor variables included factors that may influence the activity level or population density of pikas (section, time of day, site × year, and season) and factors that could influence detectability (individual observer, experience, and observer type). We log transformed hay pile volume, and square-root transformed sightings/occupied home range, calls/occupied home range, and scat/survey effort to adjust for fan-shaped residuals (i.e., heteroscedasticity, which violates assumptions of our statistical tests). Next, we used all-possible-subsets selection (Draper and Smith 1981:296–302) in R (R Core Development Team 2008) to select the suite of variables that best predicted each response (pika signs). Since individual observer and experience were collinear, we used 2 separate models to evaluate those variables. For each response variable, this procedure compares all possible combinations of the predictor variables and computes the amount of variance (R^2) explained by the model. Since we selected, a priori, no more than 5 independent variables for each sign type, we established a suitable sample size to variable ratio (Tabachnick and Fiddell 1983:92). Next, we selected the top model for each response variable and calculated the F-test P-value for each predictor variable. Finally, we determined the adjusted R^2 value for each model and the sign of correlation for each variable. To further describe the differences between detection of pika signs by professionals and volunteers, we regressed each pika sign variable by observer type (professional or volunteer) and calculated the standardized regression coefficient. The standardized regression coefficient is the difference between the means of the 2 groups divided by the residual standard error.

RESULTS

Pika Detectability Surveys Conducted by Professionals

We found evidence of observer variability during estimates of R^2 (Table 1). However, when we removed a single outlier from estimates of home ranges and calls/occupied home range, the effect of individual observer was no longer significant ($P = 0.115$ and $P = 0.209$, respectively). Time of day also influenced estimates of calls/occupied home range ($P = 0.005$). Professionals heard 38% more calls/occupied home range in the morning than during midday or evening. Professional experience level was a robust predictor of the number of scat piles found ($P < 0.001$). The most experienced professionals (level 3) found 3.8 more scat piles than professionals at level 2 and 5.7 times more than the least experienced professionals (level 1). As expected, section was the most consistent predictor for most pika signs (Table 1). Only calls/occupied home range did not differ among sections ($P = 0.188$).

Site Surveys Conducted by Volunteers and Professionals

Observer type (volunteer or professional) and site × year were the most consistent predictors of pika signs at the sites surveyed by both volunteers and professionals (Table 2). Observer type predicted estimates of occupied home ranges/survey effort ($P < 0.005$), scat/survey effort ($P < 0.001$), scat/occupied home range ($P < 0.001$), hay piles/survey effort ($P < 0.001$), hay piles/occupied home range ($P = 0.043$), and current occupants/survey effort ($P = 0.049$). In each case, professionals found more pika signs than did the volunteers (Table 3). In fact, for 5 of the 6 pika sign variables (excluding hay piles/occupied home range), the mean number found by professionals was roughly double that found by the volunteers. Professionals found 0.6–1.3 standard errors more pika signs in these 5 categories than did volunteers (Fig. 1). Observer type did not predict current occupants/occupied home range or sightings/survey effort ($P > 0.50$). Site × year was a robust estimator of all pika signs ($P < 0.05$) except scat/occupied home range ($P = 0.275$) and hay piles/survey effort ($P = 0.051$).

Table 1. Comparison of pika signs found across 8 habitat sections by each member of a trained 10-person professional crew ($n = 80$) during 2-day pika detectibility trials from 12 to 13 July 2009, in Glacier National Park, Montana, USA. Significant predictors each had an F-test $P < 0.05$, and were selected from 2 regression models for each sign using predictors: 1) section, time of day, and individual observer and 2) section, time of day, and experience. We used 2 models because individual observer and experience were collinear and thus could not be included in the same model. The adjusted R^2 of the linear regression model including the predictors listed is provided (R^2_{adj}).

Sign	Significant predictor(s)	R^2_{adj}
Home ranges	Individual observer, section	0.38
Occupied home ranges	Section	0.17
Scat	Experience, section	0.47
Hay piles	Section	0.64
Log(hay pile volume)	Section	0.58
Current occupants	Section	0.30
Sqrt(sightings/occupied home range)	Section	0.11
Sqrt(calls/occupied home range)	Time of day, individual observer	0.24

Table 2. Comparison of pika signs found by a trained professional crew and volunteers at 12 sites during 2008 and 2009 ($n = 22$), in Glacier National Park, Montana, USA. Significant predictors each had an F-test $P < 0.05$, and were selected from linear regression of the sign type against observer type, site × year, time of day, and season. The adjusted R^2 of the linear regression model including the predictors listed is provided (R^2_{adj}).

Sign	Significant predictor(s)	R^2_{adj}
Occupied home ranges/survey effort	Observer type, site × year	0.44
Sqrt(scat/survey effort)	Observer type, site × year	0.44
Scat/occupied home range	Observer type	0.21
Hay piles/survey effort	Observer type	0.28
Hay piles/occupied home range	Observer type, site × year	0.24
Current occupants/survey effort	Observer type, site × year, season	0.45
Current occupants/occupied home range	Site × year, season	0.33
Sightings/survey effort	Site × year, time of day	0.37

Season exhibited a positive correlation with 2 metrics of current occupation: current occupants/survey effort ($P = 0.017$) and current occupants/occupied home range ($P = 0.020$). Additionally, sightings/survey effort were more frequent during the evening than at midday ($P = 0.018$).

Sitting Surveys Conducted by Volunteers

Time of day was the most consistent predictor of pika signs during sitting surveys (Table 4). Volunteers observed a greater number of current occupants ($P = 0.008$), current occupants/survey effort ($P = 0.002$), and calls ($P = 0.012$) during the morning and evening than during midday. Site × year predicted current occupants/survey effort ($P = 0.019$) and calls ($P = 0.009$). Detections of calls were more frequent later in the season ($P = 0.049$).

DISCUSSION

We quantified observer variability among and between professionals and volunteers. The 2 groups observed a variety of signs that can be used to directly and indirectly identify a mammalian sentinel of climate change, the American pika. We determined that pika occupancy data collected by trained professional crews and volunteers displayed little within-group variability, but significant between-group variability. Our findings should inform managers about study design considerations to more closely align pika monitoring initiatives to their monitoring goals.

Many of the pika sign estimates varied depending on the time of day or time of year during which the observation occurred. This was particularly true for transient pika signs, such as sightings, calls, and current occupant estimates. Monitoring efforts, whether volunteer- or professional-based, should take daily activity patterns into consideration and surveys should be preferentially conducted during mornings and evenings. However, citizen science volunteers may not be as motivated to conduct surveys during hours that are outside their normal schedule. Additionally, the difficulty of accessing talus fields, often located on mountaintops or far away from trails, may make conducting surveys during ideal times of day impractical for professionals or volunteers. In this case, instructing surveyors to survey at a fixed time during midday to reduce the variability may be necessary.

The optimal monitoring season depends largely on the goals of the research effort. In Glacier National Park, snow-cover obscures many talus patches well into July. August and September surveys have dual benefits of maximum talus accessibility and high pika activity (increased hay-caching behavior). However, pika weaning often occurs in early July, so that by late August and September, many previously unoccupied and sometimes marginal territories are inhabited by young-of-the-year. This may explain our findings that more pika calls were detected and greater estimates of current occupants were recorded during this later portion of the season. Although annual turnover for pikas can average 40% (Smith and Weston 1990), death-rates for young during their first winter can be much greater and fluctuate markedly (Chris Ray, University of Colorado-Boulder, personal communication).

Table 3. Comparison of pika signs found by a trained professional crew and volunteers at 12 sites during 2008 and 2009 ($n = 22$), in Glacier National Park, Montana, USA. Survey effort was measured as person-minutes, or the number of persons surveying multiplied by the number of minutes spent surveying.

Pika signs	Professionals $\bar{x}$ (SD)	Volunteers $\bar{x}$ (SD)
Occupied home ranges/survey effort	0.119 (0.060)	0.063 (0.049)
Scat/survey effort	0.141 (0.108)	0.039 (0.063)
Scat/occupied home range	1.146 (0.696)	0.601 (0.758)
Hay piles/survey effort	0.079 (0.046)	0.040 (0.048)
Hay piles/occupied home range	0.730 (0.433)	0.687 (0.664)
Current occupants/survey effort	0.067 (0.048)	0.038 (0.043)
Current occupants/occupied home range	0.555 (0.253)	0.622 (0.420)
Calls/survey effort	0.153 (0.170)	0.032 (0.048)
Calls/occupied home range	1.377 (1.604)	0.496 (0.543)
Calls/current occupant	2.407 (2.168)	0.756 (0.552)
Sightings/survey effort	0.015 (0.017)	0.010 (0.022)
Sightings/occupied home range	0.155 (0.187)	0.177 (0.292)
Sightings/current occupant	0.258 (0.302)	0.290 (0.372)
Total signs/survey effort	0.387 (0.258)	0.122 (0.105)
Total signs/occupied home range	3.407 (2.433)	1.953 (0.916)

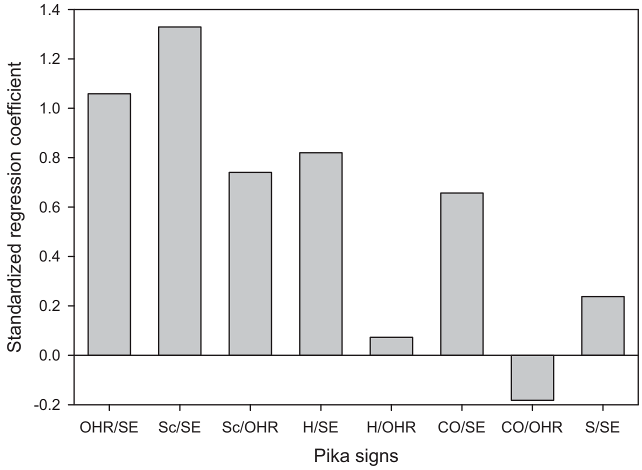

Figure 1. Comparison of pika signs found by a trained professional crew and volunteers at 12 sites during 2008 and 2009 ($n = 22$), in Glacier National Park, Montana, USA. Each pika sign variable was regressed by observer type (professional or volunteer). Standardized regression coefficient is the difference between the means of the 2 groups divided by the residual standard error. Positive values reflect a greater number of signs found by professionals. Pika signs included occupied home ranges/survey effort (OHR/SE), scat/survey effort (Sc/SE), scat/occupied home range (Sc/OHR), hay piles/survey effort (H/SE), hay piles/occupied home range (H/OHR), current occupants/survey effort (CO/SE), current occupants/occupied home range (CO/OHR), and sightings/survey effort (S/SE).

Therefore, despite greater detection rates, August and September pika densities may be an overestimation of overall fecundity and density of reproductive individuals. Surveys conducted during these months may be preferred, however, for achieving monitoring goals that rely on maximized detection probability, such as detection of large-scale occupancy of talus fields rather than in individual territories.

Professionals can consistently identify pika signs, even with minimal experience. However, we detected an individual observer effect for estimation of potential home ranges and detection of calls, and scat. We expected home range estimates to pose a consistency challenge, as numerical estimations of space and area are particularly susceptible to observer variability (Erwin 1982). Consistent estimates of potential home ranges are important for accurately determining the amount of pika habitat present and thus for monitoring efforts that measure population density, so care was taken during professional training to follow a consistent procedure.

Table 4. Significant predictors of pika signs observed by volunteers during sitting surveys at 21 sites during 2008 and 2009 ($n = 95$), in Glacier National Park, Montana, USA. Significant predictors each had an F-test $P < 0.05$, and were selected from linear regression of sign type against site × year, time of day, and season. The adjusted R^2 of the linear regression model including the predictors listed is provided (R^2_{adj}).

Sign	Significant predictor(s)	R^2_{adj}
Current occupants	Time of day	0.05
Current occupants/ survey effort	Time of day, site × year	0.26
Calls	Time of day, site × year, season	0.33

On the other hand, we were surprised to find individual observer variations in the number of calls detected during surveys by professionals. Our analyses revealed that these effects came entirely from 1 outlying observer. The individual consistently under-reported calls and may have misunderstood the instructions to record every call heard, even when coming from the same individual pika. By quantifying this observer variability, we were able to remove the data from that professional, and have a greater degree of confidence in the consistency of the remaining estimates of potential home ranges. Conducting these trials also allowed us to identify the outlier observer and work with them to mitigate the variability during future surveys.

Experience played a role in the number of scat-piles found by professionals. The effect was not attributed to a single observer and followed a predictably linear pattern of more scat piles found by the more experienced observers. In this case, the observer variability may have been influenced by the difficult terrain, as Jenkins and Manly (2008) found for ungulate fecal sampling. Traversing talus, which can be unpredictable, unstable, sharp, and precipitous, is challenging. Experienced observers typically traverse talus quicker and more efficiently, allowing them to cover more ground per unit time and still locate scat-piles that may go unnoticed by observers with less experience (L. Moyer-Horner, personal observation).

Despite effects from individual observers in the 3 areas discussed above, these analyses show remarkable consistency in the ability of professionals to detect pika signs. Our findings are consistent with other, albeit largely unpublished, reports of high detectability for pikas (Beever et al. 2003, 2010; Ray and Beever 2007; Rodhouse et al. 2010). Professionals did not exhibit observer variability when estimating the number of occupied home ranges or the number of current occupants. These are 2 of the most influential variables for determining pika presence and density. Neither experience-based variation in scat-pile identification nor underreporting of calls by an individual observer seems to have an effect on these crucial metrics.

Ideally, a professional crew would participate in detectability trials 3 times during a project: 1) at the outset, 2) midway through, and 3) at the conclusion. This would allow the researcher to obtain a comprehensive grasp on intra-crew variations and tendencies. We understand however, that measuring observer variability for each individual observer in a study is not always practical. However, these results suggest that, in general, professionals with some training and at least a few prior surveys can consistently estimate hay pile volume; the number of potential home ranges, occupied home ranges, scat piles, hay piles, and current occupants; and sightings and calls per occupied home range.

Less consistency existed between professionals and volunteers. The observer type (professional or volunteer) had significant influence on the detection of all types of pika signs except for the current occupants/occupied home range, and the number of actual pika sightings. Additionally, professionals consistently found more signs per unit of survey effort. Experience level may influence observers' ability to detect animal signs (e.g., Kendall et al. 1996). Indeed, we reported a similar effect of experience level among professionals. First-time observers, such as some of the volunteers during these surveys, may tend to identify fewer of the signs for which they are looking (Newman et al. 2003). We did not include in our analysis the experience level of each volunteer based on the number of surveys they had previously conducted. This information may be helpful in future studies to determine whether volunteers should be expected to conduct a minimum number of surveys before their data are included.

Our observation regarding experience-level and talus-traversing efficiency may also explain this discrepancy of lower detection of static signs by volunteers. In the case of pika surveys, such experience includes the physical mobility and endurance required to efficiently traverse talus sites in order to produce consistent density estimates. Professionals who are more often primed by this day-to-day field work are typically more suited to this task than volunteers who may only conduct a few surveys each season. Additionally, scat identification requires frequently bending over and looking carefully under boulders. This is physically demanding and requires a relatively high level of motivation for the observer. Volunteers may also experience a decrease in their motivation to continue searching vigorously for pika signs once a few signs have been identified. Future research could explore whether this is the case. If so, citizen science monitoring programs for pikas could be more efficient and accurate if they focused only on detection of presence–absence.

Volunteers and professionals did not differ in the number of pika sightings and estimated current occupants/occupied home range. Volunteers conducting sitting surveys also found similar numbers of current occupants per unit of survey effort at each individual site. Our research suggests that volunteers attain precise estimates of the number of current occupants, but that detections of other signs were lower than that of professionals and their estimates of density were lower relative to survey effort. For monitoring programs focused on long-term trend monitoring rather than accurately estimating density, estimates by volunteers may be adequate.

Our findings suggest that both professionals and volunteers can reliably monitor the presence of pikas at known locations, an important metric for detecting range shifts and site extirpations. We suggest that monitoring efforts with more extensive goals need to measure and address observer variability and recognize volunteer motivations and limitations to produce high-quality data (Danielsen et al. 2005, Cohn 2008). Studies that employ both groups, however, should attempt to reduce this variability through additional training for volunteers, or avoid pooling their data. Such studies should also measure and address between-group variability in their analyses. If population-level information for pikas is desired, professionals should be used to obtain density estimates. Sitting-surveys for pikas may be the best method for minimizing observer variability in monitoring programs that use volunteers. If presence–absence surveys are the focus, sitting surveys alone by volunteers should be sufficient, thus reducing the level of observer variability introduced during surveys requiring talus traversing. Thus, data collected by volunteers may inform us of possible site extirpations or colonizations, but may not provide reliable trend indicators of population size. Additionally, in some instances, volunteer surveys focused on presence–absence may inform managers on where to send professional crews.

MANAGEMENT IMPLICATIONS

American pikas are sentinels of climate change effects that can be detected by both professionals and volunteers. However, whether data are collected by the former, the latter, or a mixture of the 2, individuals should be subject to periodic observer variability trials, and managers should consider the goals of their monitoring initiative. If comprehensive population data are desired, especially those that require documentation of static pika signs such as hay piles and scat, professionals should be employed. In this case, the experience level of the professionals should be considered because the more experienced observers may find more of these static signs.

Presence–absence surveys, focused on pika sightings and estimates of numbers of current occupants, can be conducted with minimal observer variability by either group. Volunteers may be particularly appropriate for long-term monitoring in programs that receive enough initial funding to employ professionals to determine pika population status at sites, but that need to continue with decreased funding. Finally, if volunteers are used for monitoring of population trends, such as density, their training should emphasize the identification of static signs and omitting data from less experienced observers may be necessary. Managers designing such citizen science programs should consider the physical abilities of participants to determine how to offer opportunities for them to engage in the research at a higher level (e.g., talus-traversing and density estimates).

ACKNOWLEDGMENTS

We thank E. Beever for draft comments and contributions to this manuscript. We also thank A. Ives for help conceiving the idea. L. Moyer-Horner and J. Belt were supported by the Glacier National Park Fund.

LITERATURE CITED

Alldredge, M. W., K. Pacifici, T. R. Simons, and K. H. Pollock. 2008. A novel field evaluation of the effectiveness of distance sampling and double independent observer methods to estimate aural avian detection probabilities. Journal of Applied Ecology 45:1349–1356.

Beever, E. A., P. F. Brussard, and J. Berger. 2003. Patterns of apparent extirpation among isolated populations of pikas (*Ochotona princeps*) in the Great Basin. Journal of Mammalogy 84:37–54.

Beever, E. A., C. Ray, P. W. Mote, and J. L. Wilkening. 2010. Testing alternative models of climate-mediated extirpations. Ecological Applications 20:164–178.

Bonney, R., C. B. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. V. Rosenberg, and J. Shirk. 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. BioScience 59: 977–984.

Cohn, J. P. 2008. Citizen science: can volunteers do real research? BioScience 58:192–197.

Danielsen, F., N. D. Burgess, and A. Balmford. 2005. Monitoring matters: examining the potential of locally-based approaches. Biodiversity and Conservation 14:2507–2542.

Diefenbach, D. R., D. W. Brauning, and J. A. Mattice. 2003. Variability in grassland bird counts related to observer differences and species detection rates. Auk 120:1168–1179.

Draper, N., and H. Smith. 1981. Applied regression analysis. Second edition. John Wiley, New York, New York, USA.

Erwin, R. M. 1982. Observer variability in estimating numbers: an experiment. Journal of Field Ornithology 53:159–167.

Heller, N. E., and E. S. Zavaleta. 2009. Biodiversity management in the face of climate change: a review of 22 years of recommendations. Biological Conservation 142:14–32.

Inouye, D. W. 2008. Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. Ecology 89:353–362.

Jenkins, K. J., and B. F. J. Manly. 2008. A double-observer method for reducing bias in fecal pellet surveys of forest ungulates. Journal of Applied Ecology 45:1339–1348.

Kendall, W. L., B. G. Peterjohn, and J. R. Sauer. 1996. First-time observer effects in the North American Breeding Bird Survey. Auk 113:823–829.

Lindenmayer, D. B., and G. E. Likens. 2011. Direct measurement versus surrogate indicator species for evaluating environmental change and biodiversity loss. Ecosystems 14:47–59.

Lotz, A., and C. R. Allen. 2007. Observer bias in anuran call surveys. Journal of Wildlife Management 71:675–679.

MacKenzie, D. I., J. D. Nichols, J. A. Royale, K. H. Pollock, L. L. Bailey, and J. D. Hines. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier Academic Press, Burlington, Massachusetts, USA.

Newman, C., C. D. Buesching, and D. W. Macdonald. 2003. Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation. Biological Conservation 113:189–197.

Overdevest, C., C. H. Orr, and K. Stepenuck. 2004. Volunteer stream monitoring and local participation in natural resource issues. Human Ecology Review 11:177–185.

Ray, C., and E. A. Beever. 2007. Distribution and abundance of the American pika (*Ochotona princeps*) within Lava Beds National Monument. Final Report to Lava Beds National Monument, California, USA.

Rodhouse, J. R., E. A. Beever, L. K. Garrett, K. M. Irvine, M. R. Jeffress, M. Munts, and C. Ray. 2010. Distribution of American pikas in a low-elevation lava landscape: conservation implications from the range periphery. Journal of Mammalogy 91:1287–1299.

Root, T. L., J. T. Price, K. R. Hall, S. H. Schneider, C. Rosenzweig, and J. A. Pounds. 2003. Fingerprints of global warming on wild animals and plants. Nature 421:57–60.

Scott, J. M., F. W. Davis, R. G. McGhie, R. G. Wright, C. Groves, and J. Estes. 2001. Nature reserves: do they capture the full range of America's biological diversity? Ecological Applications 11:999–1007.

Silvertown, J. 2009. A new dawn for citizen science. Trends in Ecology and Evolution 24:467–471.

Smith, A. T. 1974. The distribution and dispersal of pikas: influences of behavior and climate. Ecology 55:1368–1376.

Smith, A. T., and M. L. Weston. 1990. *Ochotona princeps*. Mammalian Species 352:1–8.

Spaulding, R., P. R. Krausman, and W. B. Ballard. 2000. Observer bias and analysis of gray wolf diets from scats. Wildlife Society Bulletin 28:947–950.

Tabachnick, B. G., and L. S. Fiddell. 1983. Using multivariate statistics. Harper and Row, New York, New York, USA.

Trumbull, D. J., R. Bonney, D. Bascom, and A. Cabral. 2000. Thinking scientifically during participation in a citizen-science project. Science Education 84:265–275.

Tyser, R. W. 1980. Use of substrate for surveillance behaviors in a community of talus slope mammals. American Midland Naturalist 104:32–38.

United States Fish and Wildlife Service [USFWS]. 2010. Endangered and threatened wildlife and plants; 12-month finding on a petition to list the American pika as threatened or endangered. Department of the Interior, Washington, D.C., USA.

Associate Editor: Jeff Bowman.